

A study of genetic analysis using novel rapidly mutating Y-STR multiplex for Qatari population

Eida Almohammed ^{a,b,*}, S. Hadi^b

^aQatar Ministry of Interior-Qatar State, Doha, Qatar

^bUniversity of Central Lancashire, School of Forensics and Applied Sciences, Preston, UK

Abstract

Differentiating male lineages using non-recombining Y-chromosomal genetic markers is highly informative for tracing human migration and for forensic studies. The aim of this study was to develop allelic frequency to evaluate the resolution power of Rapidly Mutating Y-STR (RMY-STR) on Qatar population in comparison to Worldwide population. In the Qatar population the haplotype resolution of 100% was achieved. The new set of RMY-STR markers showed remarkable haplotype resolution power in the Qatar population, high gene diversity and sufficient robustness. The 13 RMY-STRs of one novel multiplex PCR reaction in the were used to type 250 samples from unrelated Qatari males. A total of 250 different haplotypes were found, with no shared haplotypes between individuals. The overall haplotype diversity (HD) was 100%. It was found that the markers which contributed the most toward high HD were DYF399S1 and DYF403S1a/b. AMOVA test performed between Qatar population in comparisons to Gulf countries, Middle East, and several worldwide population data sets were undertaken for F_{ST} values; geography was found to account considerably for the pattern of population sub structuring, although discrepancies were rarely observed in the relationship between geographical and genetic distance.

Keywords: Rapidly Mutating Y STRs Forensic Qatari population

1. Introduction

1.2 A Novel Multiplex Amplifying 13 RMY-STRs in Qatari population

Qatar occupies the land of the Qatar Peninsula, covering a surface of 11,427 km², which extends Northward from the Eastern coast of the Arabian Peninsula into the Arabic Gulf. Doha is the capital city that lies on the Eastern side of Qatar peninsula. The key position of Qatar in the Arabic Gulf was the main cause for the seasonal migrations of Arab tribes from the Arabian Peninsula and specifically from the Nejd desert (6). A novel multiplex comprising of all the 13 Rapidly Mutating Y STRs (RMY-STR) were originally described and has been developed at the University of Central Lancashire (UCLAN). The RMY-STR multiplex enabled population genotyping for Qatari population (1). The results prove that the RMY-STR multiplex is a robust and reliable amplification kit capable of overcoming inhibitors (1).

2. Materials and Methods

2.1 Buccal Samples collection

A total number of 250 saliva samples were collected from unrelated national Qatari individuals from different regions of Qatar. The samples were collected using a sterile swab that was dried and placed

in -20 °C until analysis. An informed consent was obtained for the purpose of research from all volunteers providing the samples.

2.2 Extraction, PCR Amplification & Capillary Electrophoresis

The samples were extracted using the QIAamp[®] extraction DNA Mini protocol. The extracted DNA samples were quantified using NanoDrop 2000 (Thermo Scientific Fisher) in concordance with Manufacturer instructions. A 15 µl multiplex assay amplifying 13 RMY-STR markers was used in this study. PCR were conducted in Veriti[®] 96-Well Thermal Cycle. Platinum[®] Multiplex PCR Master Mix was used in all PCR reactions using optimized thermal cycling conditions. Extraction negative, positive control 2800M (Promega), and amplification negative control were genotyped in each batch of PCR. The 13 RMY-STRs were amplified using optimized thermal PCR cycling amplified samples and electrophoresed in 8 capillary arrays on the ABI 3500 Genetic Analyzer filled with POP-6[™] polymer.

2.3 Statistical Analyses

Haplotype diversities, pairwise F_{ST} , and analysis of molecular variance (AMOVA) were calculated with Arlequin v3.5 and GenAlEx 6.41 (8). Haplotype diversities were calculated. Molecular relationships between Qatari and other Middle Eastern population samples were analysed using Network v4.6.1.1 (Fluxus Technology Ltd) applying the median-joining method. Only single-copy markers were used in the network analysis (4). The amount of genetic structure in the dataset of Qatar was evaluated by calculating F_{ST} values for the relative total population and for pairwise comparisons between regional groups.

3. Results

3.1 RM Y-STR Genotypes and Haplotype Resolution

The allele frequency and haplotype frequencies for Qatari populations were performed using Arlequin v3.5 and GenAlEx 6.41. The allele frequency of Qataris population was developed (Fig1-Supplementary Material). Microvariants were rarely observed, except in DYS393, DYS404, DYS403, DYS403a and DYS627. The gene diversity is relatively high for the Qatari population ($GD=0.901$). The highest gene diversity among loci is DYS393 for the Qatar population. The gene diversity of DYS526b fragment is slightly higher than fragment DYS526a in Qatar ($GD=0.877$). The DYS576 marker is the lowest gene diversity across all other markers Qatari populations ($GD=0.736$). The results show 100% haplotype resolution in the Qatar population. A complete (100%) haplotype resolution was achieved in the national Qatari population, and Middle Eastern group RMY-STR multiplex. In the total worldwide data, 0.9999985 haplotypes resolution was resolved with the 13 RMY-STRs. The result of network analysis of RMY-STRs of haplotypes show a visual representation of substructure with different clusters and less branching which shows less chance of relationship between haplotypes. It provides no clustering of haplotype of the two populations.

3.2 Genetic substructure analysis of Qatar population compared to other populations with RMY-STRs

Analysis of molecular variance using (AMOVA) was conducted using 1023 permutation settings for all 13 RM Y-STRs in Arlequin (8). In order to further exemplify the results, both pairwise F_{ST} and linearized Slatkin's pairwise F_{ST} resulted were plotted as matrix using R software package incorporated within Arlequin software v3.5. The genetic structure of the Qatar populations was evaluated by calculating the F_{ST} value for the total population and pairwise F_{ST} comparisons between them. Also, the comparison was performed to the other Gulf countries of (Qatar, Middle East, Kuwait, KSA, Bahrain, UAE and Yemen) when compared to the worldwide data set. The result of

pairwise F_{ST} is significant which indicate that the haplotype frequencies are different between populations. Gulf countries published data were grouped as on group region since they are all located in close geographic region after the analysis between Qatar and worldwide data. They were grouped since they are in one region in the Middle East and compared to the published Middle East data (Turkey). The results show the following values: The pairwise F_{ST} between Qatar and UAE was ($F_{ST}= 0.00957$); between Qatar and Bahrain, it was ($F_{ST}=0.01037$); between the Qatar and KSA populations; the medium F_{ST} value between Qatar and Yemen was ($F_{ST}=0.01723$); between Qatar and Middle East was ($F_{ST}=-0.01059$) and finally the average pairwise F_{ST} was ($F_{ST}= 0.00648$). The average F_{ST} value between Qatar, gulf countries and Middle Eastern population was ($F_{ST}=0.01387$). Overall, the majority of the resulted pairwise F_{ST} values are less than 0.2 which also indicate a very low substructure between populations as mentioned previously (4).

4. Discussion

The allele frequency of rapidly mutating Y-STR found in the Qatar population is not surprising and as expected. The male lineages can differ between regions that can affect the frequency distribution. Qatar population is small and shows a high degree of inbreeding where the population frequency distribution can be manipulated by genetic drift which causes the loss of alleles in comparison to the other alleles. Qatar population has been shown to have a high degree of inbreeding, so it is not surprising that a haplotype containing such a rare allele appears for this population. Marker DYS576 has high gene diversity in Qatari. The three most informative ssY-STRs DYS570, DYS576, and DYS481 are identified as three of the five Y-STRs that comprised the smallest set of loci in the nine MH Y-STRs leading to 100% resolution in six regional German, one Dutch and one Turkish population sample (3), (4).

Marker DYS399S1 in the Qatar population shows high gene diversity (0.992). The markers of DYS626, DYS526a and DYS627 have consistently high gene diversities across multiple populations, unlike many markers that show substantial variation between groups that shows a strong relationship between mutation rate and diversity of Y-STRs (7). Marker DYS449 is the most complex Y-STR with high gene diversity in Qatar population (9). Many studies have successfully genotyped DYS449 within a range of multiplexes, and this marker can successfully increase the haplotype resolution of common 9 Y-STR profiles. Multicopy markers show the highest diversity values in the Qatar population due to their highest mutation rates and the increased probability of mutation across two or more alleles within the marker.

The RMY-STR was achieved (100%) haplotype diversity resolution in the Qatari population. All the haplotypes are unique and show no sharing haplotypes between males when compared to the same geographic region in the gulf countries, Middle East and worldwide data set (2). These Gulf population are located in the same geographic area and they share the same cultural history, and same breeding mechanism or behavioural characteristics (cultural, social). The haplotype diversity from global study by Ballantyne *et al.*, (2014), was estimated at 0.99999985 with unique 12,156 haplotypes. The F_{ST} value for the Qatar population and the other populations when compared were low, which indicates relatively little population substructure suggesting the use of RMY-STR Markers. The Matrix pairwise shows no cluster found for haplotypes consisting of Qatar. The results in Qatar populations F_{ST} pairwise in comparison to the other geographical close countries Middle East population and other worldwide date showed that, the sub structuring was not an issue using RMY-STR haplotypes, on account of the high mutational rate of the markers involved. The rate of consanguinity varies within Qatar, Kuwait, UAE, Bahrain, KSA, and Middle East, the differences usually related to religion, race, ethnicity and sociocultural factors, including socially accepted norms of endogamy in tribal societies (5). In order to compensate for microsatellite distance and F_{ST} , Slatkin (1995) suggested the R_{ST} value (10). In all, this meant that (4) were unable to apply R_{ST} calculations as the application may skew the results and suggest invalid genetic structure (4). Ballantyne *et al.* (2014) demonstrate similar trends. Using F_{ST} values, most of the populations studied showed loose clustering with moderate dispersion across both dimensions. The global F_{ST} was as low at 0.00017. Between the regional groups, the average pairwise F_{ST} value was only

0.000127, with a maximum value of 0.00058 observed between Aboriginal Australians and Middle Eastern populations.

5. Conclusion

The development of an allelic frequency of a novel multiplex comprising of 13 RMY-STRs database for Qatari population will add to the scientific knowledge. This study helped to investigate the new set of RMY-STRs in the Qatari population. The outcome of this study will enhance the RMY-STR discrimination capacity for unrelated individuals in the Qatar population. It will aid to investigate the human genetic structure of Qatar population and would be able to investigate the unresolved forensic cases in a Forensic Laboratory.

6. Acknowledgement

I would like to thank all the anonymous voluntary Qatari donors of the DNA samples used in this study, without those contributions none of this works would have been possible. A special thanks to Qatar Embassy for their help and support for my project.

7. Reference

1. Alghafri, R., Goodwin, W. & Hadi, S. 2013, "Rapidly mutating Y-STRs multiplex genotyping panel to investigate UAE population", *Forensic Science International: Genetics Supplement Series*, vol. 4, no. 1, pp. e200-e201.
2. Ballantyne, K.N., Goedbloed, M., Fang, R., Schaap, O., Lao, O., Wollstein, A., Choi, Y., van Duijn, K., Vermeulen, M., Brauer, S., Decorte, R., Poetsch, M., von Wurmb-Schwark, N., de Knijff, P., Labuda, D., Vezina, H., Knoblauch, H., Lessig, R., Roewer, L., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M.R. & Kayser, M. 2010, "Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications", *American Journal of Human Genetics*, vol. 87, no. 3, pp. 341-353.
3. Ballantyne, K.N., Keerl, V., Wollstein, A., Choi, Y., Zuniga, S.B., Ralf, A., Vermeulen, M., de Knijff, P. & Kayser, M. 2012, "A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages", *Forensic Science International: Genetics*, vol. 6, no. 2, pp. 208-218.
4. Ballantyne, K.N., Ralf, A., Aboukhalid, R., Achakzai, N.M., Anjos, M.J., Ayub, et al. 2014, "Toward male individualization with rapidly mutating y-chromosomal short tandem repeats.", *Human mutation*, vol. 35, no. 8, pp. 1021-32.
5. Bener A, Hussain R, Teebi A, S, Consanguineous Marriages and Their Effects on Common Adult Diseases: Studies from an Endogamous Population. *Med Princ Pract* 2007; 16:262-267
6. CADENAS, A. M., ZHIVOTOVSKY, L. A., CAVALLI-SFORZA, L. L., UNDERHILL, P. A. & HERRERA, R. J. (2007). Y-chromosome diversity characterizes the Gulf of Oman. *European Journal Of Human Genetics*, 16 (3): 374.
7. D'Amato, M.E., Ehrenreich, L., Cloete, K., Benjeddou, M. & Davison, S. 2010, "Characterization of the highly discriminatory loci DYS449, DYS481, DYS518, DYS612, DYS626, DYS644 and DYS710", *Forensic Science International: Genetics*, vol. 4, no. 2, pp. 104-110.
8. Excoffier, L. & Lischer, H.E.L. 2010, "Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows", *Molecular Ecology Resources*, vol. 10, no. 3, pp. 564-567.

9. Kayser, M. and A. Sajantila. 2001. Mutations at Y-STR loci: Implications for paternity testing and forensic analysis. *Forensic Sci Int* 118 (2–3):116–21.
10. SLATKIN, M. 1995, "A Measure of Population Subdivision Based on Microsatellite Allele Frequencies", *Genetics*, vol. 139, no. 1, pp. 457-462.